Practical Linked Data: Learning How to Create and Use Linked Data in the Real World

Library Technology Conference, March 20, 2014
St. Paul, MN
What we talk about when we talk about linked data
A formal statement of concepts used within a knowledge domain, and the relationships between those concepts

- Genealogical ontologies
- Taxonomic ontologies ("King Philip Came Over From Greece Saturday")
- FRBR
- Dublin Core
How: The names of things, structured data, and URIs

Principles of linked data:

1. Use URIs for the names of things
2. Use http URIs so they can be looked up
3. When someone looks up a URI, provide useful information
4. Include links to other URIs

Why?
Making the Anchor Archive Subject Thesaurus available as linked data
Anchor Archive Subject Thesaurus

Developed for a special collection of zines in Halifax, Canada

Zine Thesaurus of Subject Terms

ABANDONED BUILDINGS
- RT
  - GHOST TOWNS
  - URBAN DECAY

ABLEISM
- RT
  - DISABILITIES
  - PHYSICAL DISABILITIES

ABORIGINAL PEOPLES
- USE
  - INDIGENOUS PEOPLES

Anchor Archive Zine Library
Open Metadata Registry allows people and organizations to “identify, declare or publish” metadata schemas.
Why the Open Metadata Registry?

- **Pros**
  - Creates URIs for vocabulary terms
  - Expresses vocabularies in RDF (SKOS)
  - Easy to use
  - Web based

- **Cons**
  - No bulk import (yet)
  - No spell checks
  - No de-duplication
Simple Knowledge Organization System

Main elements of the SKOS data model
Our working group rebuilt the Anchor Archive Subject Thesaurus in the Open Metadata Registry.
Step One: Entering the terms

<table>
<thead>
<tr>
<th>Detail for Alternative energy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preferred Label:</strong></td>
</tr>
<tr>
<td>Alternative energy</td>
</tr>
<tr>
<td>This is the SKOS:prefLabel property of this Concept and is required for every Concept</td>
</tr>
<tr>
<td><strong>URI:</strong></td>
</tr>
<tr>
<td><a href="http://metadataregistry.org/uri/aast/1187">http://metadataregistry.org/uri/aast/1187</a></td>
</tr>
<tr>
<td>This URI has been generated by the system, but you may override it</td>
</tr>
<tr>
<td><strong>Top Concept?:</strong></td>
</tr>
<tr>
<td>[ ]</td>
</tr>
<tr>
<td>Is this a Top Concept?</td>
</tr>
<tr>
<td><strong>Status:</strong></td>
</tr>
<tr>
<td>Published</td>
</tr>
<tr>
<td>This is the overall status of this Concept. Individual properties may have a different status</td>
</tr>
<tr>
<td><strong>Language:</strong></td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>This is the default language for this vocabulary and can't be edited here. To change the language of the prefLabel for this concept you must edit the property directly</td>
</tr>
</tbody>
</table>

- Delete
- List
- Save
- Add Concept
- Save and add
- Cancel
Step Two: Relationship mapping

**Vocabulary: Anchor Archive Subject Thesaurus**

**Concepts:** Alternative energy

<table>
<thead>
<tr>
<th>skos:property</th>
<th>Value</th>
<th>Language</th>
<th>Status</th>
<th>Updated</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>preferred label</td>
<td>Alternative energy</td>
<td>English</td>
<td>Published</td>
<td>2 February 2013 14:50</td>
<td></td>
</tr>
<tr>
<td>has broader</td>
<td>Energy</td>
<td></td>
<td>Published</td>
<td>17 February 2013 20:25</td>
<td></td>
</tr>
<tr>
<td>related to</td>
<td>Energy efficiency</td>
<td></td>
<td>Published</td>
<td>17 February 2013 20:26</td>
<td></td>
</tr>
<tr>
<td>related to</td>
<td>Climate change</td>
<td></td>
<td>Published</td>
<td>5 June 2013 17:08</td>
<td></td>
</tr>
<tr>
<td>has narrower</td>
<td>Biodiesel</td>
<td></td>
<td>Published</td>
<td>5 June 2013 17:09</td>
<td></td>
</tr>
</tbody>
</table>

5 results

[Add Property]
**Completed term in Open Metadata Registry**

**Vocabulary:** Anchor Archive Subject Thesaurus  
**Concepts:** Alternative energy

<table>
<thead>
<tr>
<th>Metadata +</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Detail</strong></td>
</tr>
<tr>
<td>Preferred Label:</td>
</tr>
<tr>
<td>Language:</td>
</tr>
<tr>
<td>URI:</td>
</tr>
<tr>
<td>Top Concept?:</td>
</tr>
<tr>
<td>Status:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>related to</td>
</tr>
<tr>
<td>related to</td>
</tr>
<tr>
<td>preferred label</td>
</tr>
<tr>
<td>has narrower</td>
</tr>
<tr>
<td>has broader</td>
</tr>
</tbody>
</table>
<!-- Concept: Alternative energy -->

<skos:Concept rdf:about="http://metadataregistry.org/uri/aast/1187" xml:lang="en">
  <skos:inScheme rdf:resource="http://metadataregistry.org/uri/aast"/>
  <reg:identifier rdf:resource="5170"/>
  <skos:prefLabel xml:lang="en">Alternative energy</skos:prefLabel>
  <skos:related rdf:resource="http://metadataregistry.org/uri/aast/1214"/>
</skos:Concept>
Open Metadata Registry provides an RDF/XML serialization of the SKOS framework for the whole vocabulary (or any individual term).

This is a SKOS file that you can download and save somewhere.
Twincity Zine Linked Data Working Group

Linking the Anchor Archive Subject Thesaurus to the MCTC Zine Collection
Zines: a self publication, motivated by a desire for self expression, not for profit.
Open Refine

- Open Refine is a free open source tool for cleaning up large metadata sets. It can also be used to create linkages between metadata sets.

**Free Your Metadata**
Learn how to get more value out of metadata easily

- [freeyourmetadata.org](http://freeyourmetadata.org) has instructions for using Open Refine to clean up metadata and link it to other data sets.
Using Open Refine to Link

Step 1: Place your metadata in Open Refine.

Step 2: Make sure you have the RDF extension.
Using Open Refine to Link

Step 3: Using the RDF extension, you set up a reconciliation service.

• I will need to use based on RDF file....
Using Open Refine to Link

- Name the reconciliation service
- Upload your SKOS file from Open Metadata Registry
- Make sure you check SKOS in the label properties area
### Using Open Refine to Link

**Step 4:** Select the column you want to reconcile, and navigate to *start reconciling*...

- I want to match my Subject Headings
Using Open Refine to Link

Step 5: Select your newly created reconciliation service

- In my case, it’s AAST (Anchor Archive Subject Thesaurus)
Using Open Refine to Link

Step 6: Open Refine will reconcile your metadata with the linked data of your choice.

- If you have lots of metadata, this is the “Go get some Coffee” step.
Using Open Refine to Link

Step 7: Select the best matches

- The blue headings are linked to the Anchor Archive Subject Thesaurus in the Open Metadata Registry
Using Open Refine to Link

- You can match a single subject heading at a time
Using Open Refine to Link

- Or you can match multiple similar cells at once
Thoughts on the Linked Data Project

• MCTC Zine Collection metadata is not ready for prime time

• Open Metadata Registry can also be linked through a SPARQL endpoint

• MCTC Library has had major cuts to staff this year
What Can We Do Now?

Some of the uses to which it could be put
Basically what we’ve done is create URIs for these subject headings, giving them a permanent central location on the web.
But that doesn’t really mean anything unless someone actually uses the subject headings within a collection.
One thing we could do is enter a collection of zines, such as the one at MCTC, into an online database and use the URIs for the subjects that we entered into the Open Metadata Registry.
One database possibility to explore is Freebase.com, "a community-curated database of well-known people, places and things."
“Freebase is a database of things”
“All things in Freebase have strong identifiers”
The data in Freebase is harvested from certain websites (like Wikipedia) or is directly entered by individuals like you or me.
Things in Freebase can also be entered in RDF form, which means that the content in Freebase can be used as linked data.
If we chose to, we could take MCTC’s zine collection and enter it in Freebase.

In the metadata we create for each zine in Freebase, we could link to the subject URIs that we created in the Open Metadata Registry.
Each zine entered in Freebase would become linked data that could be used on other websites. Each zine would also link back (through the subjects) to more zines on the same topic.
Or at least, that’s our theory... 😊
We could use the vocabulary data to provide other kinds of support for searching

• Non-preferred terms

• Broader, narrower, and related terms

• Scope notes

• Synonym rings
Another possibility would be to map terms in our Zine vocabulary to other vocabularies.
Public transit
(http://metadataregistry.org/uri/aast1012)

Skos:closeMatch

Local transit
(http://id.loc.gov/authorities/subjects/sh85077945)
Publication of ISO 25964-2
Information and Documentation—Thesauri and Interoperability with Other Vocabularies
Part 2: Interoperability with Other Vocabularies

MARC Discussion Paper No. 2014-DP02
Relationships between subject headings from different thesauri in the MARC 21 Authority Format
Another way to let searchers connect with our resources
ID: 300311452

**cinnabar (mineral and pigment)** (mineral, inorganic material, ... Material

**Note:** A soft, dense, red, native ore composed of mercuric sulfide, found in deposits at springs around the world. Cinnabar was mined in the Roman Empire for use as a pigment, and has been the main source of mercury throughout the centuries. The toxic stone was used as a talisman in Mayan tombs, as a colorant in Chinese carved lacquerware, in ink used by the Incan Empire, and other applications. Cinnabar is chemically identical to synthetic red vermillion.

**Terms:**

- **cinnabar (mineral and pigment)** *(preferred, C, U, English-P, D, U, N)*
- **cenobrium** *(C, U, English, UF, U, N)*
- **natural vermilion** *(C, U, English, UF, U, N)*
- **cinabre** *(C, U, French-P, D, U, N)*
- **cinabrio** *(C, U, Spanish-P, D, U, N)*
- **cinabro** *(C, U, Italian-P, D, U, N)*
- **cínabrio** *(C, U, Portuguese-P, UF, U, N)*
Publishing Authorities and Vocabularies

LC Linked Data Service
Authorities and Vocabularies

Search

Enter Keyword or Phrase

Available Datasets

The Linked Data Service provides access to commonly found standards and vocabularies promulgated by the Library of Congress. This includes data values and the controlled vocabularies that house them. The following are currently offered as part of this service:

- LC Subject Headings
- LC Name Authority File
- LC Classification
- LC Children's Subject Headings
- MARC Relators
- MARC Countries
- MARC Geographic Areas
- MARC Languages
- ISO639-1 Languages
- ISO639-2 Languages
- ISO639-3 Languages
- Extended Date/Time Format
Project: Location LODer

Location LODer

Logainm.ie

Altna Bulloch Atha Clith - Naomh Dublin
Type: city County: Dublin

DBpedia

Dublin is the capital and most populous city of Ireland. The English name for the city is derived from the Irish name Dubhlinn, meaning “black pool”. Dublin is situated near the midpoint of Ireland’s east coast, at the mouth of the River Liffey and the centre of the Dublin Region. Originally founded as a Viking settlement, it evolved into the Kingdom of Dublin and became the island’s principal city following the Norman invasion. The city expanded rapidly from the 14th century, it was briefly the second largest city in the British Empire and the fifth largest in Europe. Dublin entered a period of stagnation following the Act of Union of 1800, but it remained the economic centre for most of the island.

National Library of Ireland

1. A map of a piece of ground on the north side of Parliament Street in the City of Dublin belonging to Deed, Filton. Surveyed by J. Longfield. April 1824
2. A map of that part of the Royal Canal intended to pass through the County of Dublin from Summer Hill to Ash-town, by John Browning 1790. Scale 500 feet to an inch. Table of reference showing demesnes Proprietors Leases under-Tenants & remarks on the properties
3. A map of several holdings on West, Arran St, Phoenix St, & Bow Lane in the City of Dublin, surveyed for Francis Prendergast by J.L. 1613. Belonging to Lord Clare
4. A survey of lands and premises lying between Ashmun St and Prussia St done for General Hawkins, by John Longfield 1813. Scale 40 feet to an Inch. Tenants names show on

Europeana

1. A smocky chimney
2. Tram in Dublin
3. Blue Tit
4. Coal Train
5. Common Frog
6. Grey Squirrel
7. Hen
8. Irish hare
9. Red Fox
10. Robin on a rock in the Phoenix Park Dublin
11. Construction
12. Dublin Docklands Development
13. Grand Canal Square
15. United Nations Secretary General Ban Ki-moon and Her Majesty Queen Margrethe II of
Project: Linked Jazz
What we learned